Refining the conformational ensembles of flexible proteins using simulation-guided spectroscopy

Jennifer M. Hays

Blue Waters Symposium
June 3 2018
Acknowledgements

• Kasson Lab
  • Peter Kasson, MD, PhD
  • Ricardo Ferriera, PhD
  • Eric Irrgang, PhD
  • Kenta Okamoto, PhD
  • Ana Pabis, PhD
  • Anjali Senger, PhD
  • Cara Broshkevitch

• Columbus Lab
  • Linda Columbus, PhD
  • Marissa Kieber, PhD

[Logos for University of Virginia, ARCS, Blue Waters, and NIH]
Flexible proteins are hard to study because many states contribute to their conformational ensembles.
Iterative refinement leverages both computation and experiment.

Identify informative experiments

Perform experiments

Integrate experimental data using computation

\[ U_{bias} = \frac{1}{2} k (P_{MD}(x) - P_{DEER}(x))^2 \]
Opa-CEACAM interaction

CEACAM1
N-terminal Ig Domain
PDB ID: 2GK2

How we study binding: DEER

• Double electron-electron resonance (DEER) provides distance distributions between pairs of spin-labeled amino acids

• Great! Captures conformational heterogeneity of Opa ensemble

• **Problem:** DEER only reports on a few DOF in the system, but we need on the order of ~10,000 measurements
mRMR Algorithm
(maximum-relevancy, minimum redundancy)

We want two things from our pairs:

1. maximally informative on the conformational ensemble
2. are minimally redundant with each other

Hays et al., Submitted
Integration of spectroscopic data

After 100 ns of simulation per ensemble member (2 μs aggregate), approximately converge to target

$$U = \frac{1}{2} K (P_{MD}(x) - P_{DEER}(x))^2$$

Hays et al., Submitted
As a teaser...

Current method not robust enough to handle well-separated probability modes

Come see my poster tonight for more details!
How well-refined are the resulting ensembles?

Hays et al., Submitted
Two rounds of simulation-guided spectroscopy generate a hypothesis for CEACAM engagement

Opa loops have three regions of high variability: SV (tan), HV1 (green), HV2 (red)

SV extended: 40 %

HV2 extended: 20 %

Multiple loop extensions: 40 %
HV2-extended conformations are selected during CEACAM engagement.

Hays et al., Submitted
Summary

• By incorporating optimally-selected DEER pairs into MD simulation, we have obtained a conformational ensemble of apo Opa that suggests clear possible modes of CEACAM engagement.

• Further refinement of the bound ensemble
  • binding experiments to determine which loop(s) CEACAM binds: HV2 or the HV1/SV interface?
  • additional simulations + CEACAM

Incorporate distributions into simulations

Identify highly informative pairs

Restrained-ensemble simulations

DEER experiments

\[ U_{bias} = \frac{1}{2} k (P_{MD}(x) - P_{DEER}(x))^2 \]
Thank you
Questions?
Additional Slides
High conformational heterogeneity pairs

- Opa 31-166: entropy = 4.64 nats
- Opa 88-162: entropy = 4.71 nats

Low conformational heterogeneity pairs

- Opa 77-107: entropy = 4.45 nats
- Opa 107-117: entropy = 4.34 nats

DEER experiments

Restrained-ensemble simulations

Incorporate distributions into simulations

Identify highly informative pairs
mRMR: 2nd round
mRMR: 1st round
SSP

Dimensionality of conformational ensemble

Information-theoretic coarseness (1-ε)

Coarse → Fine

x10^{-5}